

# Linked Data in Production

Moving Beyond Ontologies

David Newbury

Assistant Director, Software and User Experience, Getty

CNI Spring Project Briefings, March 26, 2024

Thank you for inviting me to present today.

I am a **technologist** working in the cultural heritage sector.

I lead the **public digital** team, developing applications to support Getty's mission.

Introduction: **Getty**

---

I work for **Getty** in Los Angeles, which is a library/archive/museum/research center.

One of our major areas of digital leadership is in the use of **Linked Data** for **cultural heritage**.

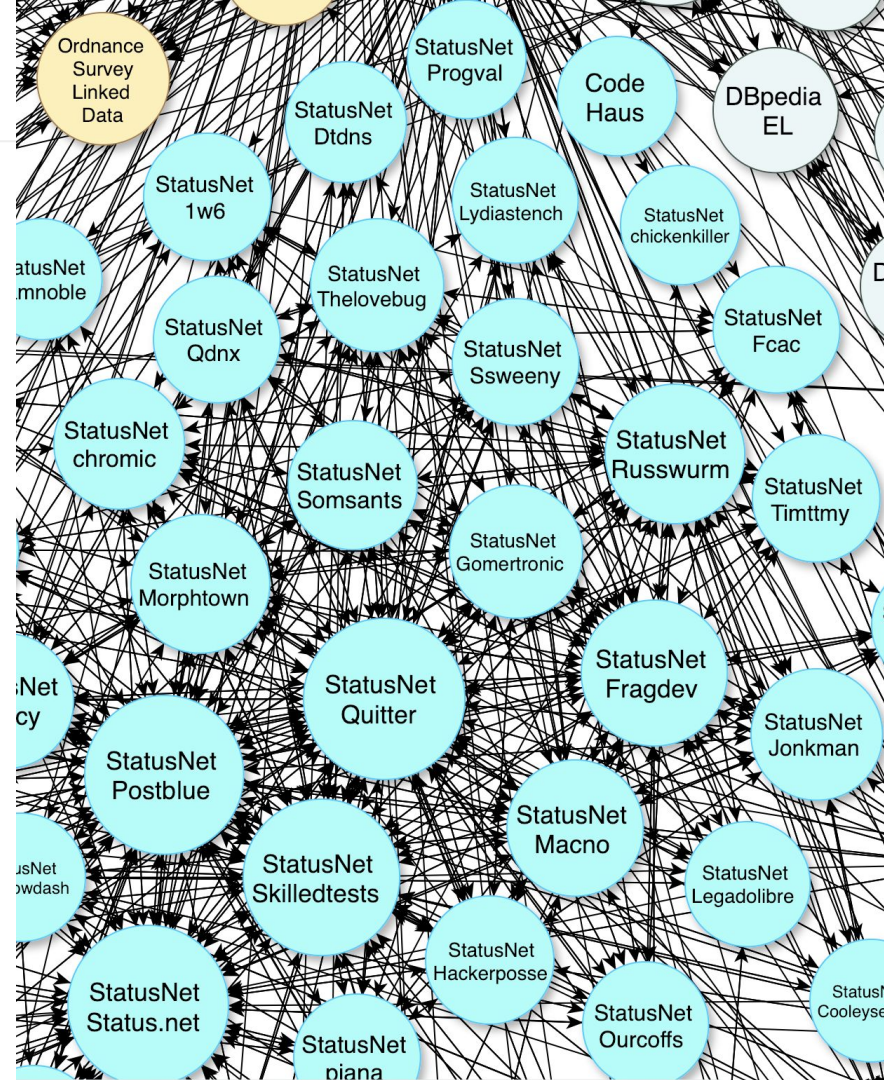
**Getty**



## Introduction: What is Linked Open Data?

Linked Open Data is a set of technologies that attempt to translate some of the best practices of the Web for use with structured data:

- The use of URLs as identifiers
- Networks of information, not tables
- Formal, shared standards for description



Introduction: **Linked Data is (mostly) Dead.**

---

As it turns out, Linked Data is not wildly successful.

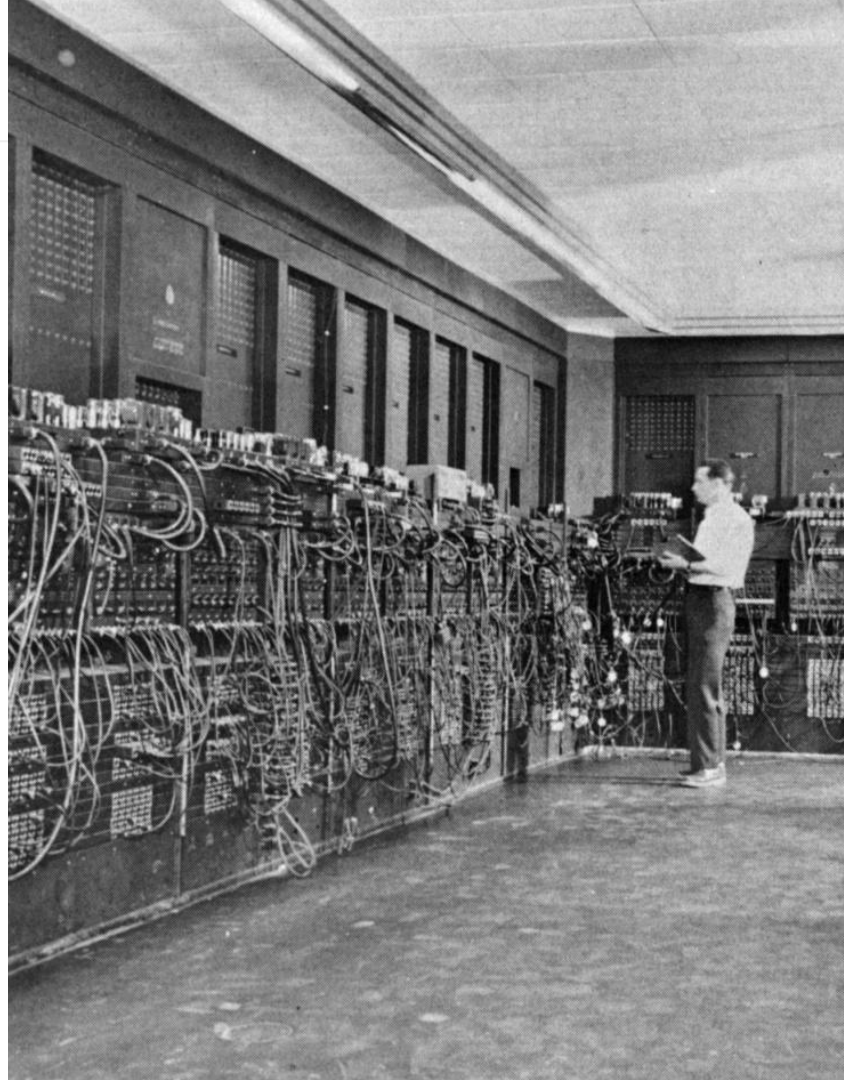
But we all still talk about it a lot.

## Introduction: **Provocation**

---

Linked Data's appeal in cultural heritage is a technological solution to a social problem:

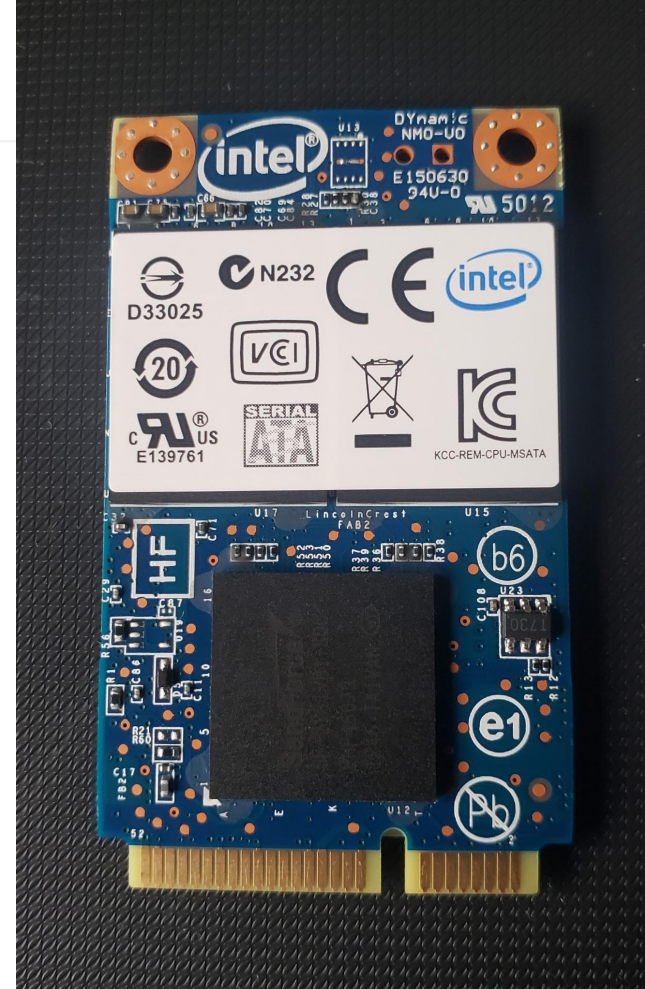
Cheap storage, ubiquitous connectivity, and search algorithms **recontextualize the labor** behind cultural heritage data work.



## Introduction: Cheap Storage

Mass digitization, computational metadata generation, and decades of cataloging mean that our institutions have more data to provide than we have the **ability to provide context** for.

Data overload and limited user attention are the collections access problems of the next decade.



# Introduction: Ubiquitous Connectivity

Our always-connected culture means that our collections are increasingly seen as part of a **single digital ecosystem**—

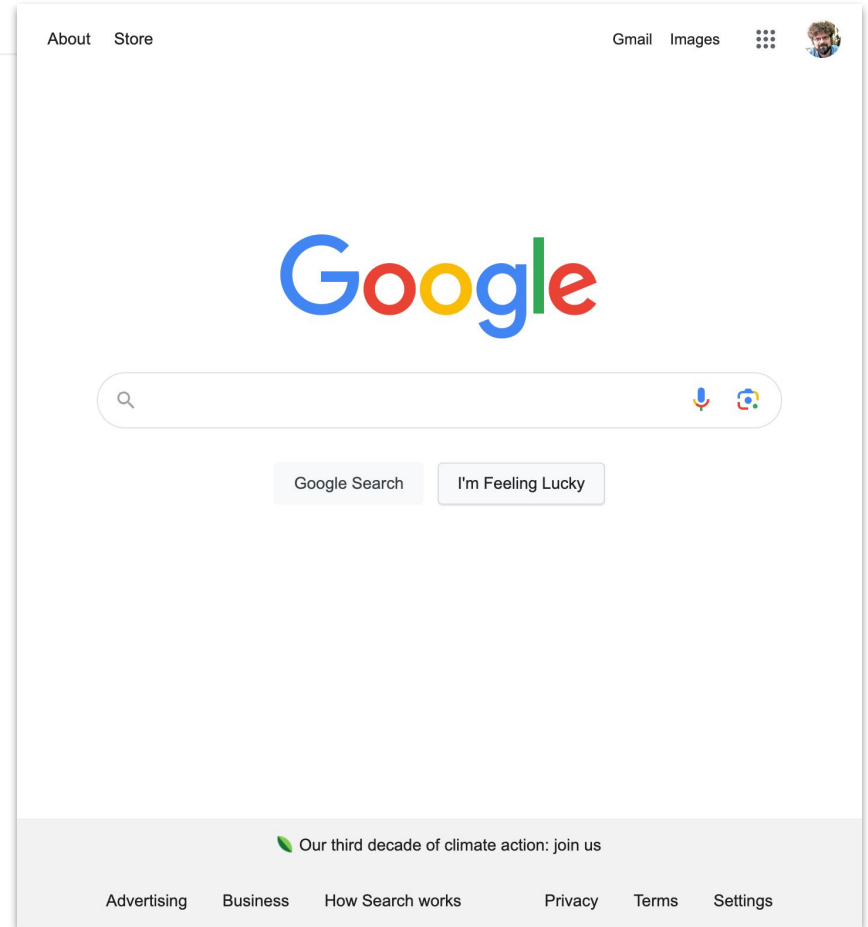
And the questions that are being asked require information that **extends beyond** the boundaries of any one institution.





And commercial tools have given users the expectation that information is available for the asking—

Eliding the labor and capital needed to create, curate, and maintain that information.



## Linked Data has been seen as a solution:

It provide structures that manage the **scale of data** we create,  
identifiers that maintain **authority** in a globally distributed environment,  
and ontologies that enable complex **data retrieval** across datasets.

What came before:

# Laying the Groundwork

# Laying the Groundwork: Getty Vocabularies

In 2014, the **Getty Vocabularies** were launched as Linked Data.

This, alongside the work at Yale Center for British Art, Rijksmuseum, and the British Museum, demonstrated the **feasibility** of LOD within the museum community.

Build 2022-09-20T05:25:06, Revision 3

## Laying the Groundwork: Carnegie Museum Archives

The archives of the Carnegie Museum of Art's Film Department launched in 2014.

The animating question was:

What would happen if you treated the **relationships** between events, archival material, people, and artwork as the essential element, **not the objects?**



VISIT EXHIBITIONS + EVENTS COLLECTION CARNEGIE MUSEUM OF ART LEARN INTERACT BLOG SHOP

cmoa archives MUSEUM ARCHIVES VISITING THE ARCHIVES HELP  
SEARCH FINDING AIDS TERMS OF USE

### Carolee Schneemann

No biographical information available.

Birth: 1939  
Place of Birth: Fox Chase, Pennsylvania  
Nationality: American

#### ARTWORKS IN THE PERMANENT COLLECTION

No Image Available  
Fuses, 1964-1967, 1964-1967

#### ARCHIVAL MATERIALS

Created by Carolee Schneemann


## Laying the Groundwork: American Art Collaborative

In 2017, the American Art Collaborative launched.

It used these same principles to highlight connections across 14 institutions and 152,000 items—using Getty’s Vocabularies as a bridging structure between institutions.



AAC Demonstration Prototype About the AAC Notice

# AMERICAN ART COLLABORATIVE

**WELCOME!**

The American Art Collaborative (AAC) is a consortium of 14 art museums in the United States. The Collaborative believes that Linked Open Data offers rich potential to increase the understanding of art by expanding access to cultural holdings, by deepening research connections for scholars and curators, and by creating public interfaces for students, teachers, and museum visitors.

This simple demonstration application lets you browse among the institutions' data that was provided for the AAC.

[Read more about the American Art Collaborative](#)

**ABOUT THE DATA**

Each institution has provided this publicly-available data and images under their own particular licenses. The licenses are not the same for all institutions! The information is available on the [Data and Image Licensing](#) page on this site.







**14**  
Institutions

**152,447**  
Items

**25,702**  
Creators

**21,660**  
Related Parties

### PARTNER INSTITUTIONS

 The Amon Carter	6,328 Items	472 Creators	1754 — Date Range
 Archives of American Art	15,771 Items	4,261 Creators	1743 — 2203 Date Range
 Autry Museum of the American West	193 Items	114 Creators	1610 — 2015 Date Range
<input type="checkbox"/> Colby College Museum of Art	8,281 Items	1,747 Creators	1100 — 2016 Date Range
<input type="checkbox"/> Crystal Bridges Museum of American Art	1,556 Items	381 Creators	1670 — 2015 Date Range
<input type="checkbox"/> Dallas Museum of Art	2,229 Items	648 Creators	1680 — 2011 Date Range
 Gilcrease Museum	20,886 Items	1,048 Creators	1000 — 19104 Date Range
 Indianapolis Museum of Art at Newfields	11,095 Items	2,147 Creators	1540 — 2014 Date Range
 National Museum of Wildlife Art	2,212 Items	375 Creators	1747 — 2104 Date Range

## Laying the Groundwork: **Linked Art**

One of the most lasting outcomes of the American Art Collaborative was **Linked.Art**, the shared data model that connected institutions.

### Digital Objects

All Digital Objects share some basic characteristics, regardless of their particular nature. The [basic patterns](#) of Name, Identifier, Classification, and Statement all apply in the regular way. Beyond the baseline, digital objects can have the following descriptive features:

- Access Point - The URLs where the object is available. These may also be treated as Digital Objects in their own right, but typically only the URL is given.
- Format - The `format` of a digital object is its media type, often called a MIME type, given as a string
- Standard - Many digital objects further conform to standard specifications, which can be referenced using the `conforms_to` property. This differs from `format`, as there may not be a media type for the specification, and from `classified_as` which is a broader classification (image, rather than conforming to the standard for JPEG 2000)
- Dimensions - Digital dimensions follow the same pattern as [physical dimensions](#), but might use different types (file size) or the same (height, width for images) and different units (bytes, pixels).
- Creation - Digital Objects are created by `Creation` events rather than `Production` events, but otherwise have the same activity model.

#### Example:

A web page was created by a Museum Education department using the HTML format (and standard), is 100k in size, and is available on the museum website.

```
{
  "@context": "https://linked.art/ns/v1/linked-art.json",
  "id": "https://linked.art/example/digital/1",
  "type": "DigitalObject",
  "_label": "Digital Object",
  "classified_as": [
    {
      "id": "http://vocab.getty.edu/aat/300264578",
      "type": "Type",
      "_label": "Web Page"
    }
  ],
  "identified_by": [
    {
      "type": "Name",
      "content": "Informative Web Page"
    }
  ],
  "dimension": [
    {
      "type": "Dimension",
      "_label": "100 kb",
      "classified_as": [
        {
          "id": "http://vocab.getty.edu/aat/300265863",
          "type": "Type",
          "_label": "File Size"
        }
      ],
      "value": 100,
      "unit": {
        "id": "http://vocab.getty.edu/aat/300265870",
        "type": "MeasurementUnit",
        "_label": "kilobytes"
      }
    }
  ]
}
```

What we've done:

# Getty's Digital Ecosystem



# Getty's Linked Data: Getty Vocabularies

Getty has been doing Linked Data since 2014, starting with the Getty Vocabularies.

It's a thesaurus of concepts, people, and places used for cataloging across many institutions.

Getty Vocabularies: LOD SPARQL Queries

Try Search Search Brief

### Tupperware Home Parties

Source: <http://vocab.getty.edu/ulan/500355972>

Subject (45) Predicate Object All Website Hierarchy Download in: JSON JSONLD RDF N3/Turtle N-Triples

Inference Explicit only

Statements in which the resource exists as a subject.

Predicate	Object
rdf:type	gvp:GroupConcept
rdfs:seeAlso	<a href="http://www.getty.edu/vow/ULANFullDisplay?find=&amp;role=&amp;nation=&amp;subjectid=500355972">http://www.getty.edu/vow/ULANFullDisplay?find=&amp;role=&amp;nation=&amp;subjectid=500355972</a>
dcterms:created	2013-11-19T14:58:18
dcterms:modified	2013-11-19T14:58:18, 2013-11-19T15:12:30, 2013-11-19T15:15:58, 2013-11-19T16:21:04, 2013-11-19T16:21:25, 2013-11-19T16:21:35, 2013-11-19T16:22:13
skos:changeNote	ulan_rev:5506508654, ulan_rev:5506508655, ulan_rev:5506508674, ulan_rev:5506508676, ulan_rev:5506508677, ulan_rev:5506508679, ulan_rev:5506508690, ulan_rev:5506508691, ulan_rev:5506508811, ulan_rev:5506508812, ulan_rev:5506508813, ulan_rev:5506508814, ulan_rev:5506508815, ulan_rev:5506508818
gvp:parentString	Corporate Bodies
gvp:parentStringAbbrev	Corporate Bodies
gvp:displayOrder	38563
skos:exactMatch	<a href="http://id.loc.gov/authorities/names/n95068543">http://id.loc.gov/authorities/names/n95068543</a>
xl:prefLabel	ulan_term:1500814043
xl:altLabel	ulan_term:1500814044
gvp:broaderInstantial	ulan:500000003
gvp:broaderPreferred	ulan:500000003
gvp:prefLabelGVP	ulan_term:1500814043
skos:inScheme	ulan
skos:scopeNote	ulan_scopeNote:123928
dcterms:contributor	ulan_contrib:2500000013
dcterms:source	ulan_source:2100149014-subject-500355972
gvp:agentTypePreferred	aat:300025969
dc:identifier	500355972
gvp:ulan1003_associated_wit...	ulan:500330449
cc:license	<a href="http://opendatacommons.org/licenses/by/1.0/">http://opendatacommons.org/licenses/by/1.0/</a>
dcterms:license	<a href="http://opendatacommons.org/licenses/by/1.0/">http://opendatacommons.org/licenses/by/1.0/</a>
void:inDataset	<a href="http://vocab.getty.edu/dataset/ulan">http://vocab.getty.edu/dataset/ulan</a>
prov:wasGeneratedBy	ulan_rev:5506508655
foaf:focus	ulan:500355972-agent

Build 2022-09-20T05:25:06, Revision 3

# Getty's Linked Data: Archival Records

Since then, we've moved most of our major systems to use Linked Data—including our archives...

The screenshot displays the Getty Research Collections Viewer interface. At the top, the navigation bar includes the Getty logo and links for Museum, Research Institute, Conservation Institute, Foundation, and Support Us. Below this, the page title is "Research Collections Viewer". A large image of a lion sculpture is shown at the top. Below the image, there are thumbnails for various items: Recto, Verso, Letter, Envelope, and another Recto. The main content area features a section titled "Sylvia Sleight correspondence with Lawrence Alloway, 1951-1990". This section includes a "Location in the Collection" tree diagram, a "Request from the Library Catalog" section, and "Digital Use Restrictions". The "Full Record Details" section provides metadata for the collection, including Title, Creator, and Classifications. The "Details From Collection" section includes Collection and Publication Rights information.

Getty

Museum Research Institute Conservation Institute Foundation Support Us

Visit What's On Explore Art Resources Our Work About

Research Collections Viewer

Home Search About Help Feedback

This object has 15 images

Recto Verso Letter Envelope Recto

**Sylvia Sleight correspondence with Lawrence Alloway, 1951-1990**

Location in the Collection

- Sylvia Sleight papers, 1803-2013, undated (bulk 1940-2000)
- Correspondence, 1874-2010, undated (bulk 1940-1990)
- General correspondence, circa 1920-2010
- Alloway, Lawrence, 1949-1990, undated
- 1951-1990
- Sylvia Sleight correspondence with Lawrence Alloway, 1951-1990 (This object)

Request from the Library Catalog

Request this item?

Box & Folder: Box 1 Folder 8

Digital Use Restrictions

Digital images and files saved from this website are for study purposes only. Copyright restrictions apply. Copyright © Paul Getty Trust. For more information, see the [Library Reproductions & Permissions](#) page. The Getty Research Institute, Los Angeles.

Full Record Details

Title: Sylvia Sleight correspondence with Lawrence Alloway, 1951-1990

Creator: Alloway, Lawrence, 1926-1990

Classifications: personal correspondence

Details From Collection

Collection: Sylvia Sleight papers

Publication Rights: Contact [Reproductions & Permissions](#)

Actions

View TIFF manifest

Open in comparison viewer

# Getty's Linked Data: Archival Records

Since then, we've moved most of our major systems to use Linked Data—including our archives...

... and our museum collection.


Museum Research Institute Conservation Institute Foundation Support Us  
Visit What's On Explore Art Resources Our Work About

## Museum Collection

Home Search Feedback

### Plastic Vase in the Form of a Female Bust

about 580 B.C.  
Unknown artist/maker  
On view at Getty Villa, Gallery 304, Archaic and Classical Greece  
[View full record details](#)



This small aryballos (oil flask) takes the shape of the bust of a woman. Her head is topped by a small neck with a disk-shaped mouth decorated with black dots placed in between four eyes. Plastic vases—usually made in the form of a human, animal, or mythical creature—were especially popular in the Greek world from about 600 to 500 B.C. They typically held perfume oil, and their narrow spouted openings were designed to conserve this precious commodity. Rhodes was a leading producer of such vessels. Busts of women were one of the most common forms of Rhodian plastic vases. This woman's facial features—large eyes, long nose, broad forehead—and breast correspond with the Rhodian sculptural style of the time. The woman wears a red chiton and a mantle over her right shoulder, as well as a beaded necklace, earrings and a hairband. Slightly modeled and spongy-gate texture to the marble. The vessel was made using a two-piece mold, into which thin sheets of clay were pressed. When slightly dried, the two halves of the vessel were joined with a clay paste and the vase was decorated.

[Order this image](#) [Image Copyright / Image: Bruce White Photography](#)

#### More from the Collection

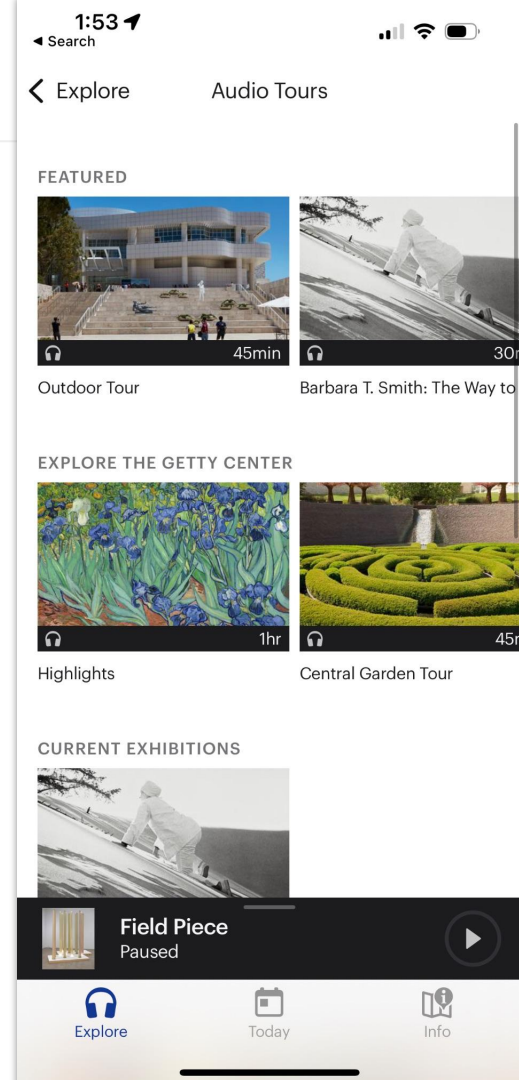
- Double-Headed Aryballos**  
Unknown
- Attic Red-Figure Neck Amphora Fragment**  
Attributed to Berlin Painter (Greek [Attic], active about 500 - about 480 B.C.)
- Attic Black-Figure Band Cup Fragment**  
Unknown
- Lilly Pheen, with wife, Richmond, Va.**  
(Edward and Henry J. Anthony & Co. (American, founded 1862, merged 1902)
- Portrait of Bess Cohen in Spanish Shawl / Bess Cohen, Spanish Shawl**  
Paul R. Haviland (American, 1890 - 1903)
- Wife on Onyx**  
William Eglington (American, born 1939)

#### Full Artwork Details

<b>Title:</b>	Plastic Vase in the Form of a Female Bust
<b>Artist/Maker:</b>	Unknown
<b>Date:</b>	about 580 B.C.
<b>Medium:</b>	Terracotta
<b>Dimensions:</b>	10.3 x 6.3 x 5.9 cm (4 1/8 x 2 1/16 x 2 5/16 in.)
<b>Place:</b>	Rhodes, Greece (Place Created)
<b>Culture:</b>	East Greek (Rhodian)
<b>Object Number:</b>	96.AE.90
<b>Credit Line:</b>	Gift of Barbara and Lawrence Fleischman
<b>Alternate Titles:</b>	Oil Container in the Form of a Bust of a Woman (Alternate Title)
<b>Department:</b>	Antiquities
<b>Classification:</b>	Vessel

## Getty's Linked Data: **Audio Guide**

It's used for onsite visitor experiences via our audio guide...



# Getty's Linked Data: 12 Sunsets

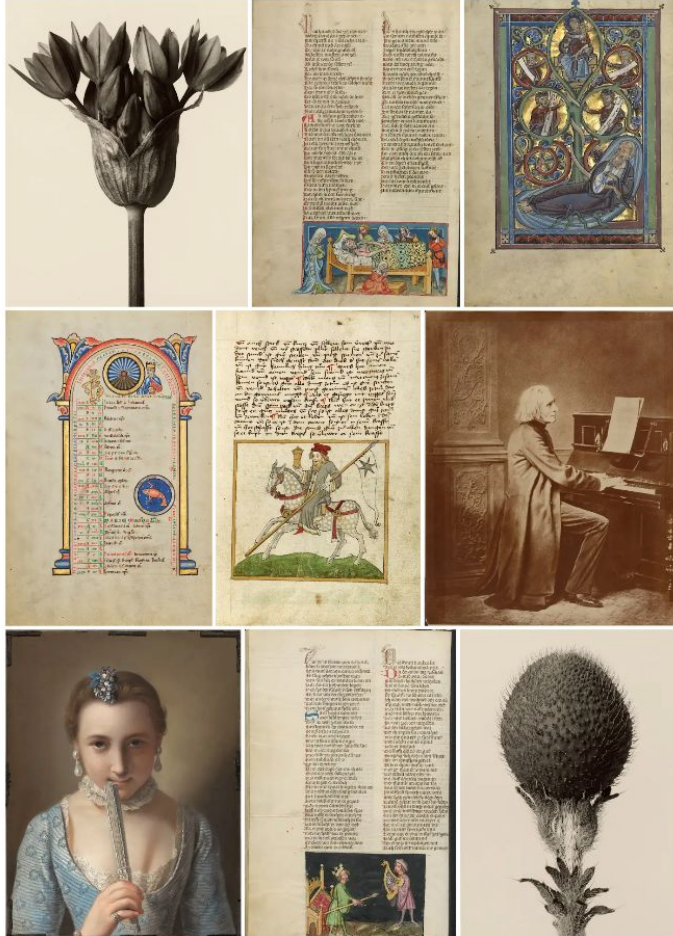
It's used for onsite visitor experiences via our audio guide...

...and to provide novel interfaces for exploration of our materials.

The screenshot displays a digital interface for '12 SUNSETS', exploring Ed Ruscha's archive. At the top, a red banner reads '12 SUNSETS' in white, followed by 'EXPLORING ED RUSCHA'S ARCHIVE'. Navigation options include 'DRIVE', 'SORT', 'ABOUT', and 'SHARE'. A 'LOCATION SEARCH' field with a magnifying glass icon is on the right, and a yellow question mark icon is in the top right corner. The main content is divided into three horizontal sections: 1) A street view image of a building with a large sign that says 'TOWER RECORDS'. A '1985' label is in the top left, and 'ADD YEAR +' is in the bottom left. A 'FLIP' button with a double-headed arrow is in the bottom right. 2) A map view showing a red truck on a yellow path between 'NORTH DOHENY DRIVE' and 'HOLLOWAY DRIVE'. The map includes 'mapbox' branding and copyright information: '© Mapbox © OpenStreetMap Improve this map'. A vertical yellow bar on the right says 'CHOOSE RIDE'. 3) An interior view of a record store with shelves of records and a sign that says 'L.A. RECORDS'. A 'Getty' logo is in the bottom left, and 'MAP' and 'FOCUS' buttons are in the bottom right. At the very bottom, a copyright notice reads: '© J. Paul Getty Trust. | Privacy Policy | Terms of Use | Trademarks | All images © Ed Ruscha. Used with permission.'

# Getty's Linked Data: Google Arts & Culture

It's also used by third parties: both large, like Google Arts & Culture...



## Getty's Linked Data: Spanish Art in the US

It's also used by third parties: both large, like Google Arts & Culture...

...and small, like this project by the Cultural Office of the Embassy of Spain.

Getty

HOME MUSEUMS

### J. PAUL GETTY MUSEUM

9.7.25 THE GETTY CENTER, 1200 GETTY CENTER DR., LOS ANGELES, CA 90049



Getty

The J. Paul Getty Museum, with locations in Malibu (Getty Villa) and Los Angeles (Getty Center), caters to diverse audiences with a wide range of art exhibitions and programs in visual arts.

The Getty Center features works of art dating from the eighth through the twenty-first century, showcased against a backdrop of dramatic architecture, terraced gardens, and panoramic views of Los Angeles. The collection includes European paintings, drawings, sculpture, illuminated manuscripts, decorative arts, and European, Asian, and American photographs.

The Getty Villa in Malibu features Greek, Roman, and Etruscan antiquities presented in a setting modeled after a first-century Roman country house, the Villa dei Papiri in Herculaneum, Italy.

[VISIT THE MUSEUM'S WEBSITE](#)

### Spanish collection

The following 34 Spanish artworks are a selection from the collection of the J. Paul Getty Museum.



### Leaf from *Commentarius in Apocalypsim*

BY SAINT BEATUS OF LIÉBANA, CIRCA 1220–1235

Medium  
Tempera colors, gold leaf, and ink

Dimensions  
Leaf: 29.4 × 23.5 cm (11 5/16 × 9 1/4 in.)

13TH CENTURY

# Getty's Linked Data: APIs

We've also built a complex, powerful digital infrastructure to support this work—millions of records in a single shared data model, pulling from a wide collection of systems of record.



Museum Research Institute Conservation Institute Foundation Support Us

Visit - What's On - Explore ART Resources Our Work About

## API Documentation

### Getty Museum Collection

#### Records about the artwork held at the Getty Museum

This Museum Collection API provides access to the metadata about the more than 290,000 objects within the Getty Museum Collection. It also contains information about the people, groups, places, events, and other supplementary information that you'd need to know in order to make sense of those records.

Last updated: 03/30/2022  
See revision history

---

#### On this page

- Getting Started
- Tracking Changes
- Asking Questions
- Usage Guidelines
- Licensing
- Exception #1: Images
- Exception #2: Written Descriptions
- Attribution
- Distant Integrity
- No Endorsements/Representation
- Questions?

---

This API Documentation is a work in progress.

With some exceptions, the data available from this API is licensed under the [CC0 1.0 Universal license](#). Check the [Usage Guidelines](#) section of this document for more details.

#### Quick Reference

This API is based on the [Linked Art](#) standard, and uses [ActivityStreams](#) to track changes and [IIF](#) to provide images. We also provide a [SPARQL](#) endpoint to allow for graph queries.

- **REST endpoints:** <https://data.getty.edu/museum/collection/> `ENTITY_TYPE=IDENTITY_ID`
- The available entry types are: `OBJECT`, `PLACE`, `DOCUMENT`, `GROUP`, `PERSON`, `EXHIBITION`, and `ACTIVITY`.
- **IIF Image API:** <https://media.getty.edu/iif/images/> `IMAGE_ID`
- **IIF Presentation API:** <https://media.getty.edu/iif/manifests/> `MANIFEST_ID`
- **ActivityStream:** <https://data.getty.edu/iif/museum/collection/activity-stream>
- **SPARQL:** <https://data.getty.edu/museum/collection/sparql>

There is also a GUI for SPARQL queries available at <https://data.getty.edu/museum/collection/sparql-ui>

---

#### Getting Started

This API is currently designed to accommodate three tasks: [getting records](#), [tracking changes](#) in records, and [asking questions](#) about the collection as a whole.

We currently don't provide a way to get a list of all of the objects or other entity types in the collection. We know how useful this would be, and it's on our roadmap. We also don't provide a way to download all the data in the collection. Also on the roadmap. Stay tuned.

#### Getting Records and Images

The first task is to access information about the records Getty has published the "entities" in our collection.

#### Records

The most obvious record type is for `OBJECT` entities, usually artwork like Van Gogh's *Innes*. However, Vincent van Gogh himself is a `PERSON` entity, and the gallery we display *Innes* in is a `PLACE` entity. Each entity within our dataset has a URL where you can access the record as a JSON document, and the JSON contains links to other entities across our collection.

For example:

- *Innes* is available at <https://data.getty.edu/museum/collection/object/088c3d70-d691-4f5b-9de7-7d76b6d4d84b>
- Vincent van Gogh is available at <https://data.getty.edu/museum/collection/person/3a876a9-5333-40d5-8488-6cc722068f6e>
- Wae Ywllon, Gallery 204 is available at <https://data.getty.edu/museum/collection/place/5339e663-6c89-411c-b34c-7d32cc62706e>

You can learn more about the model we use on the [linked art](#) website, and about our particular implementation details in the [IIF Reference](#) section below.

#### Images

The other kind of record we provide are images of the artwork. We provide these via our Getty-wide [IIF API](#)—a standard used across cultural heritage for accessing pictures. There are two APIs available for images—one, the IIF image API, lets you get access to `PIC`s of the images at different file sizes. The other, the IIF Presentation API, lets you get access to all the images associated with an object, presented with some contextual information, provided as a JSON document known as a IIF Manifest.

You can get access to these IIF API URLs from each Object entity record—assuming we have images for that artwork!

For example:

- The main image of *Innes* is available at [https://media.getty.edu/iif/images/65428650-7f98-4697-9540-5d4b9690d4f4/iif/500\\_0564e6a1e09](https://media.getty.edu/iif/images/65428650-7f98-4697-9540-5d4b9690d4f4/iif/500_0564e6a1e09)
- A thumbnail of *Innes* is available at [https://media.getty.edu/iif/images/65428650-7f98-4697-9540-5d4b9690d4f4/iif/500\\_0564e6a1e09](https://media.getty.edu/iif/images/65428650-7f98-4697-9540-5d4b9690d4f4/iif/500_0564e6a1e09)
- The IIF Manifest for *Innes* is available at <https://media.getty.edu/iif/manifest/5339e663-6c89-411c-b34c-7d32cc62706e>

For more information about what IIF can do, check out the API documentation on the [IIF website](#):

- [IIF Image API 2.1.1 Documentation](#)
- [IIF Presentation API 2.1.1 Documentation](#)



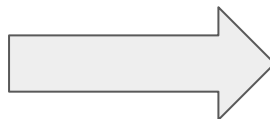
Under the hood:

# Linked Data & the Everything API

# Data Flow: How we wish it was

The screenshot shows the ArchivesSpace staff interface. At the top, it says 'ArchivesSpace a community server by eGangas'. Below that, there are navigation options like 'Browse', 'Create', and a search bar. The main content area is titled 'Krispy Kreme corporate records' and has an 'Accession' button. Under 'Basic Information', there are fields for Title, Identifier (with dropdowns for NMAH, AC, OSH, and 1), Accession Date (1987-07-17), and a 'Public?' checkbox. There are also sections for Content Description, Condition Description, Disposition, Inventory, Provenance, Retention Rule, General Note, Acquisition Type (Deposit), Resource Type (Records), Language of Description, and Restrictions (Apply?). A 'Save Accession' button is at the bottom left.

Staff Interface



The screenshot shows the Getty Research Collections Viewer public website. At the top, it says 'Getty' and has navigation links like 'Home', 'Search', 'About', 'Help', and 'Feedback'. The main content area features a large image of a lion sculpture. Below the image, it says 'This object has 15 images' and shows a grid of thumbnails for 'Recto', 'Verso', 'Letter', 'Envelope', and 'Recto'. At the bottom, there is a section titled 'Location in the Collection' with a list of related items, including 'Sylvia Sleigh papers, 1803-2013, updated (bulk 1940-2000)', 'Correspondence, 1874-2010, undated, bulk 1940-1990', 'General correspondence, circa 1920-2010', 'Alloway, Lawrence, 1849-1990, undated', and '1951-1990'. A 'Request from the Library Catalog' link is also present.

Public Website

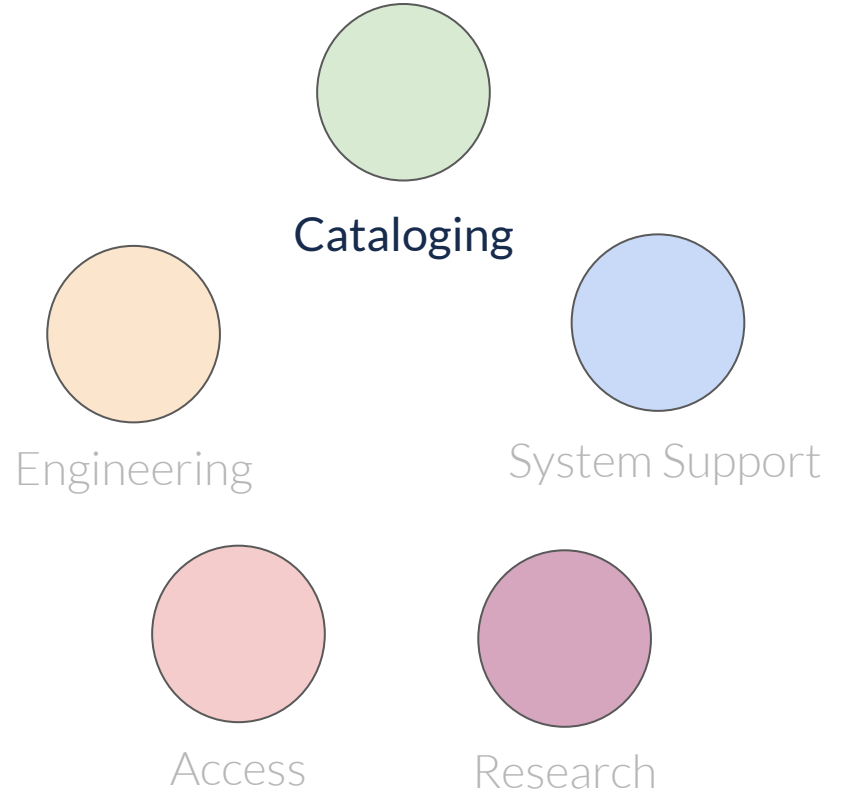


**These systems support people.**

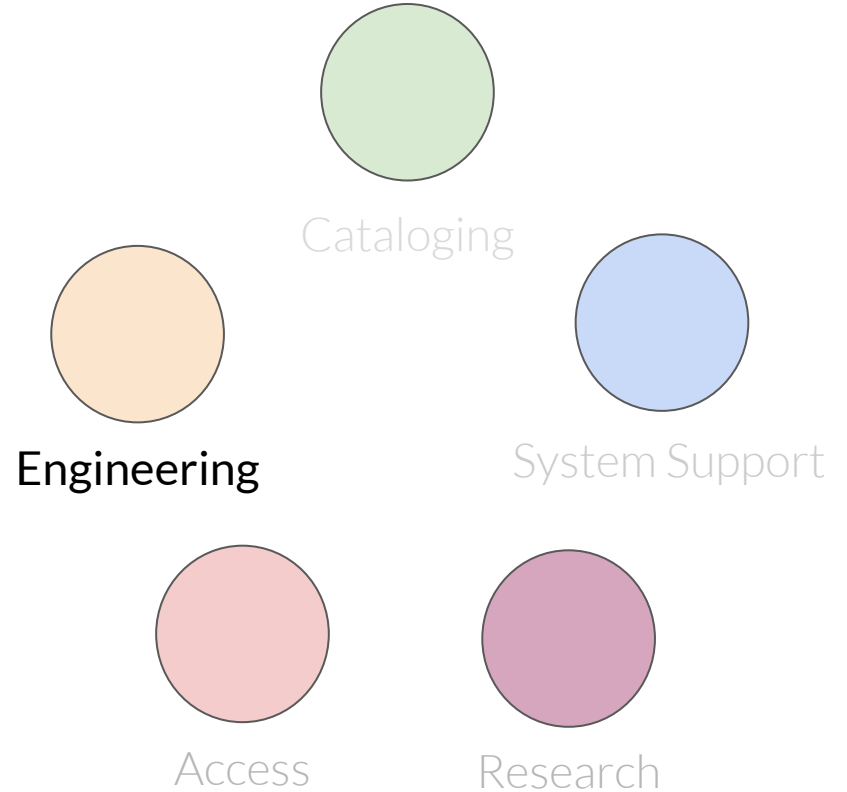
Digital infrastructure is designed to use computers to empower **people** to be more effective at meeting the mission of the organization.

Catalogers need systems that match their workflows—and different disciplines have different needs.

Our infrastructure needed to not be tied to any particular backend system.

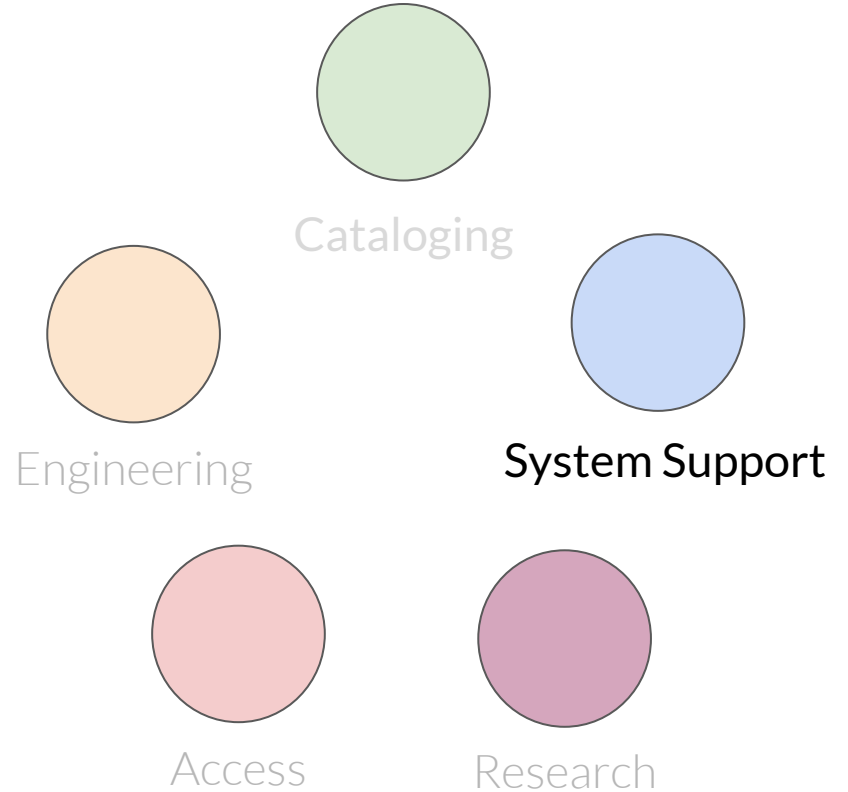


Engineers need to get data in and out of systems, using patterns and practices that they already know how to use.



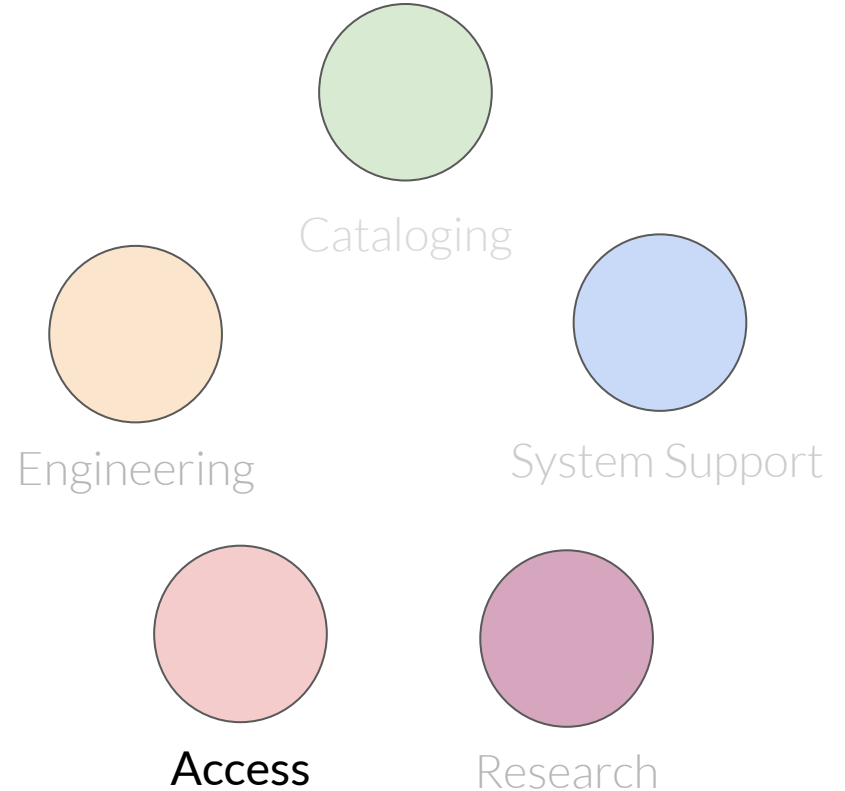
System admins just don't want you to break their stuff.

Pulling data out of systems on demand usually breaks stuff.



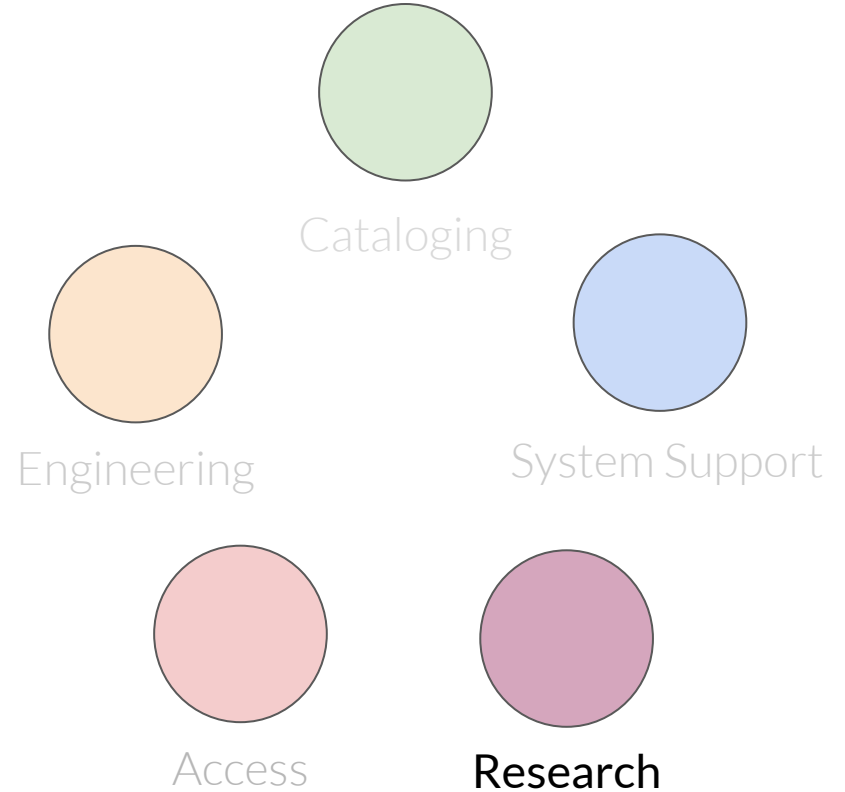
Most end users are looking for content—they want to learn what we know on a given topic.

This may be professional scholarship or it might be looking for pictures—both are examples of information-seeking behaviours.





And some researchers want to find questions that haven't been asked before—to find new connections or patterns in the data that others have overlooked.



## Getty's Linked Data: The LOD Gateway

Meeting the needs of catalogers is *mostly* not my problem.

There are high-quality, professional tools that work within the disciplinary training of the field.



The screenshot displays the ArchivesSpace web interface. At the top, the ArchivesSpace logo is visible, along with navigation options like 'Select Repository', 'System', and a user profile for 'admin'. The main header includes 'Home / Accessions / Krispy Kreme corporate records / Edit'. A left sidebar contains a menu of navigation options such as 'Basic Information', 'Languages', 'Dates', 'Extents', 'Agent Links', 'Related Resources', 'Component Links', 'Related Accessions', 'Subjects', 'External Documents', 'Rights Statements', 'Metadata Rights Declarations', 'Instances', 'Deaccessions', 'Collection Management', 'Classifications', and 'User Defined'. The main content area is titled 'Krispy Kreme corporate records' and features a 'Basic Information' section. This section includes fields for 'Title' (Krispy Kreme corporate records), 'Identifier' (NMAH, AC, 0594, 1), 'Accession Date' (1997-07-17), 'Publish?' (checkbox), 'Content Description' (Contains correspondence, administrative records, operational records, company newsletters, news clippings, photographs, photograph albums, and audiovisual materials created and compiled by the Krispy Kreme corporation between), 'Condition Description' (Generally good, the collection contains photographs that required rehousing and audiovisual material that requires reformatting), 'Disposition', 'Inventory', 'Provenance' (This collection was donated to the National Museum of American History, Archives Center on July 17, 1997, by the Krispy Kreme Corporation), 'Retention Rule', 'General Note', 'Acquisition Type' (Deposit), 'Resource Type' (Records), 'Language of Description', 'Script of Description', and 'Restrictions Apply?' (checkbox).

# Getty's Linked Data: The LOD Gateway

Providing access to that data, though, often requires recontextualization:

Changing the conceptual lens from one focused on staff efficiencies to one focused on user's needs.



## Research Collections Viewer

Home Search About Help Feedback



This object has 15 images

1 Recto 2 Verso 3 Letter 4 Envelope 5 Recto 6 Verso

### Sylvia Sleigh correspondence with Lawrence Alloway, 1951-1990

- Location in the Collection
- Sylvia Sleigh papers, 1803-2013, undated (bulk 1940-2000)
  - Correspondence, 1874-2010, undated, bulk 1940-1990
  - General correspondence, circa 1920-2010
  - Alloway, Lawrence, 1949-1990, undated
  - 1951-1990
  - Sylvia Sleigh correspondence with Lawrence Alloway, 1951-1990 (This object)

Request from the Library Catalog

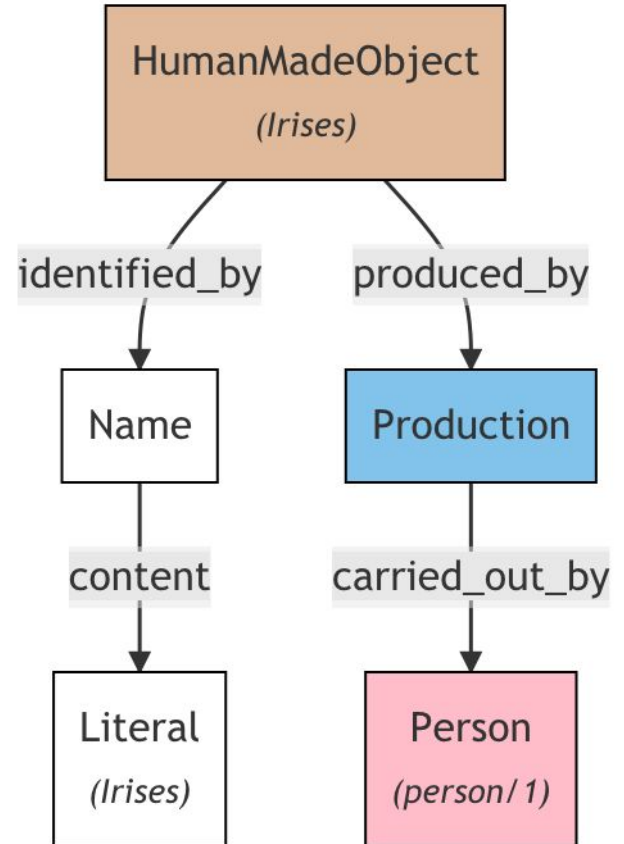
# Getty's Linked Data: The LOD Gateway

Doing so requires combining data from multiple systems and multiple workflows into a new record.

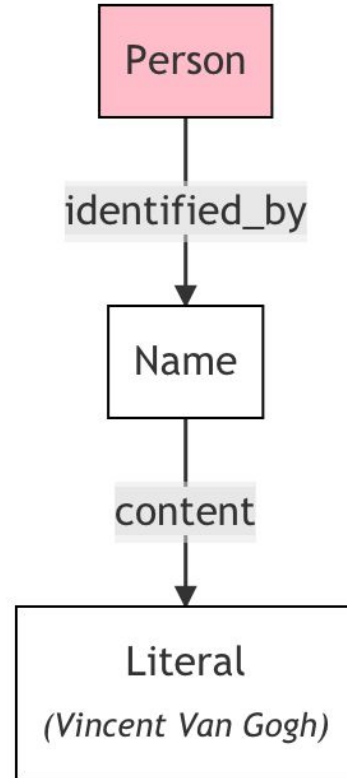
This combining—or linking—of data has tradeoffs.



Imagine a record for the painting *Irises*.



And a second record, this one for **Van Gogh**.

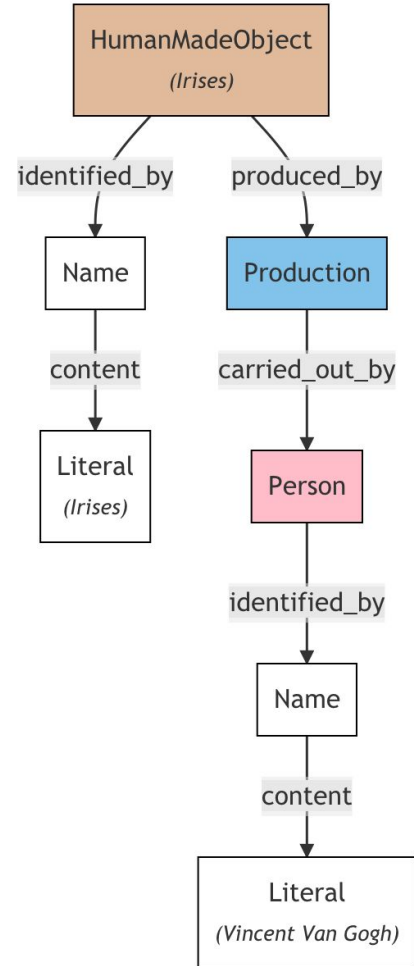


These could be seen as two separate documents:

```
"@context":  
"https://linked.art/ns/v1/linked-art.json",  
"id": "person/1",  
"type": "Person",  
"identified_by": {  
  "id": "person/1/name",  
  "type": "Name",  
  "content": "Vincent Van Gogh"  
}
```

```
"@context":  
"https://linked.art/ns/v1/linked-art.json",  
"id": "object/1",  
"type": "HumanMadeObject",  
"identified_by": {  
  "id": "object/1/name",  
  "type": "Name",  
  "content": "Irises"  
},  
"produced_by": {  
  "id": "object/1/production",  
  "carried_out_by": {"id": "person/1"}  
}
```

Or as a single graph.

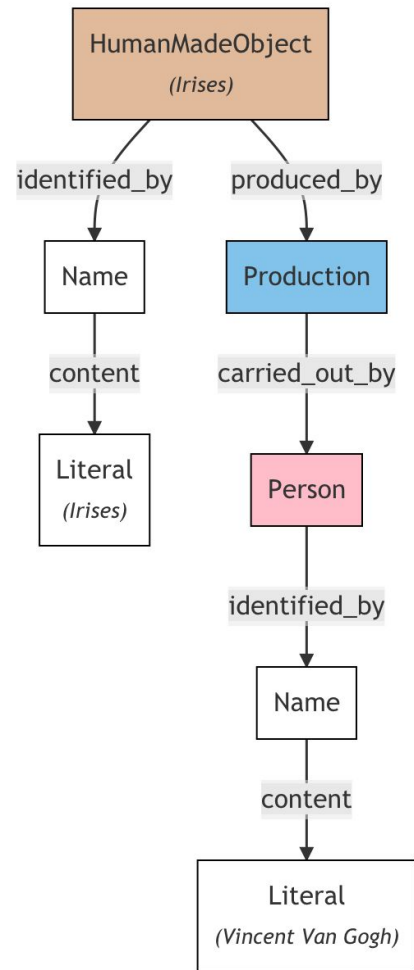




## Getty's Linked Data: The LOD Gateway

From the point of view of the data, these are equivalent—they contain the same facts.

But from a usability perspective, they make different things easy or hard.



Documents: **For Access and Discovery**

Documents are optimized for **Access:**

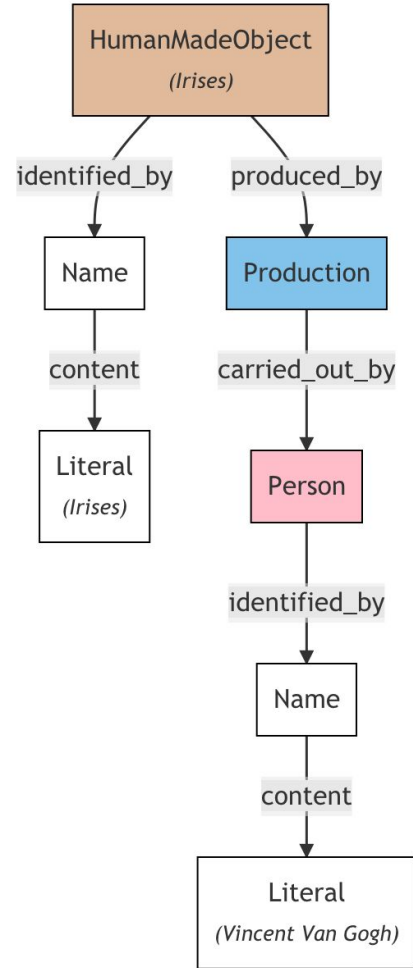
They provide a specific set of data bundled together by the data creator that provide all the facts you need...given a **specific context**.

```
"@context":  
  "https://linked.art/ns/v1/linked-art.json",  
  "id": "object/1",  
  "type": "HumanMadeObject",  
  "identified_by": {  
    "id": "object/1/name",  
    "type": "Name",  
    "content": "Irises"  
  },  
  "produced_by": {  
    "id": "object/1/production",  
    "carried_out_by": {"id": "person/1"}  
  }  
}
```

Graphs: **For Queries**

Graphs, alternately, are optimized for **querying**:

Allowing a user to define a specific context based on novel criteria and returning that subset of facts.



Imagine two Questions:

---

“What objects does Getty have that have images larger than 1200px on the longest side that have been exhibited in both New York and Paris and were created by artists who lived before 1850?

and

What’s the label info for *Iris*es?

Imagine two Questions:

---

At the Getty, we have *never* asked:

**“What objects does Getty have that have images larger than 1200px on the longest side that have been exhibited in both New York and Paris and were created by artists who lived before 1850?”**

but we ask

**What’s the label info for *Iris*es?**

Several thousand times a day.

Documents: **For Access and Discovery**

Having an interface for documents lets us provide a simple, easily understandable record that maps well to known contexts.

This is important, because people usually *expect* these contexts. It makes answering common questions simple.

## Museum Collection

[Home](#) [Search](#) [Feedback](#)

90.PA.20

### Iris

1889

[Vincent van Gogh](#) (Dutch, 1853 - 1890)

On view at  
[Getty Center, Museum West Pavilion, Gallery W204](#)

[View full record details](#)


In May 1889, after episodes of self-mutilation and hospitalization, Vincent van Gogh chose to enter an asylum in Saint-Rémy, France. There, in the last year before his death, he created almost 130 paintings. Within the first week, he began *Iris*, working from nature in the asylum's garden. The cropped composition, divided into broad areas of vivid color with monumental



## Documents: For Access and Discovery


Documents are also the way **the internet** works: REST APIs, cache control, JSON, webpages.

Using these well-known systems helps **developers** make systems that are fast and easy to build.

**Irises** 


Vincent van Gogh

---

**API URL:** [object/c88b3df0-de91-4f5b-a9ef-7b2b9a6d8abb](https://www.getty.edu/art/collection/object/c88b3df0-de91-4f5b-a9ef-7b2b9a6d8abb) 


The entity identifier in the [Museum Collection API](#)

---

**TMS ID:** 947 



The identifier in the Museum's internal collection management system


---

**DOR ID:** 826 

The identifier in the Museum's legacy internal digital object repository


---

**IIIF:**  [Manifest URL](#) 

[Open in comparison viewer](#) 


The International Image Interoperability Framework, or IIIF, is an open standard for delivering high-quality, attributed digital objects online at scale. Visit [iiif.io](#) to learn more.

---

**SPARQL API:** [Open in Query Builder](#) 


The SPARQL API provides an endpoint and interface for constructing graph queries of the Museum Collection Linked Open Data

---

**Related IDs:** [Internal Identifier Relationships](#) 

The ID management system records relationships between identifiers across Getty's systems and resources

---

**Permalink:** <https://www.getty.edu/art/collection/object/103JNH> 

The permanent url for this page in Getty's Collection Online

Last updated: August 21, 2023  
*This date refers to the last time the primary JSON-LD document that contains information on this record has changed.*

## Graphs: For Asking Questions

Research is different—each scholar brings their own question and their own context.

Meeting their need means empowering them to draw **their own** boundaries within the data.



Cole et al. Boundary contours

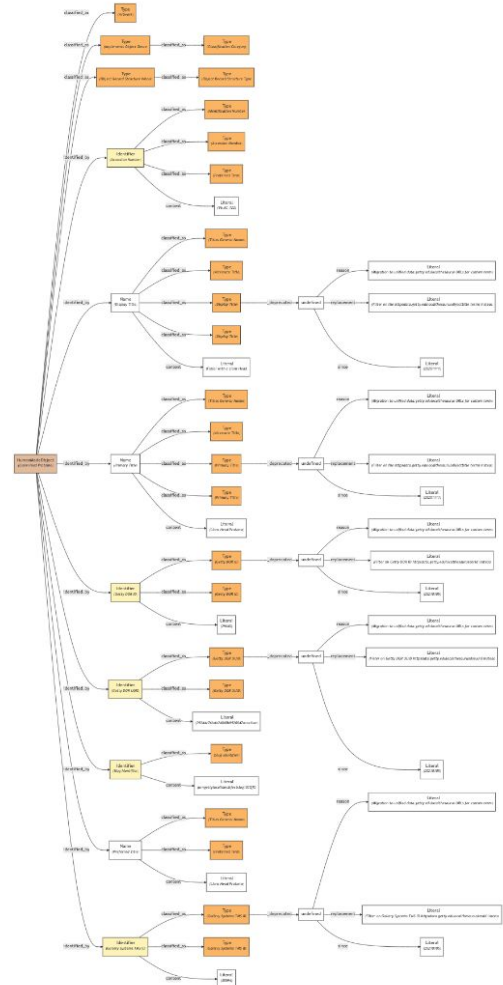
Photog par Bisson Fr. Imp. Lemercier Paris



# Graphs: For Asking Questions

Doing so is complex—it moves the burden of defining the relevant context to the **user** of the data, not the **creator** of the data.

But it makes asking new questions **possible**, even if it might be inefficient or complicated.



## Meeting Both Needs

We've built our infrastructure to allow for both use cases:

A developer can create, update, and delete documents, and behind the scenes it will keep a graph in sync with those changes.

```
Query 1
1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>
4 SELECT ?title ?artist_name WHERE {
5   ?object
6   crm:P1_is_identified_by / crm:P190_has_symbolic_content ?
  title;
7   crm:P108i_was_produced_by / crm:P14_carried_out_by /
  crm:P1_is_identified_by / crm:P190_has_symbolic_content ?
  artist_name.
8 }
```

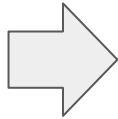
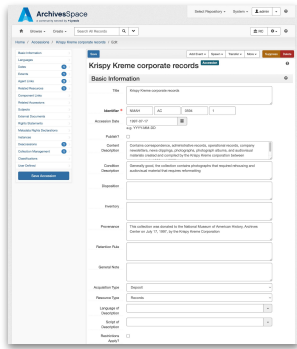
```
3 results in 0.038 seconds
Simple view  Ellipse 
title
1 Irises
2 The Night Watch
3 Starry Night
```

```
{
  @context: "https://linked.art/ns/v1/linked-art.json",
  - classified_as: [
    - {
      _label: "Subject Heading - Topical",
      id: "https://data.getty.edu/local/thesaurus/asspace-subject-topical",
      type: "Type"
    }
  ],
  content: "Art, American -- California -- 20th century",
  format: "text/html",
  id: "https://data.getty.edu/research/collections/document/54f867ec-14a2-5502-8167-bd7074f",
  - identified_by: [
    - {
      - classified_as: [
        - {
          _label: "generated URL slug",
          id: "https://data.getty.edu/local/thesaurus/temporary-slug",
          type: "Type"
        },
        - {
          _label: "Slug Identifier",
          id: "https://data.getty.edu/local/thesaurus/slug-identifier",
          type: "Type"
        }
      ],
      content: "10SF6V",
      id: "https://data.getty.edu/research/collections/document/54f867ec-14a2-5502-8167",
      type: "Identifier"
    }
  ],
  - part: [
    - {
      id: "https://data.getty.edu/research/collections/document/e6737072-d0cb-5621-a751",
    }
  ],
  type: "LinguisticObject"
}
```

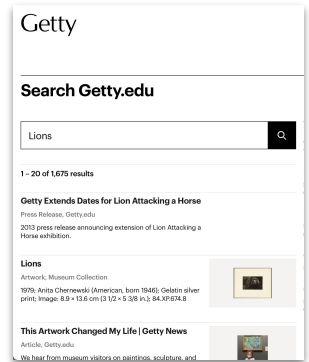
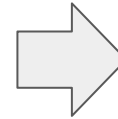
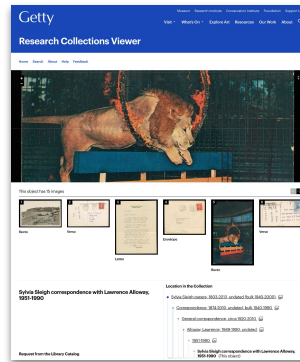
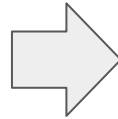
# Linked Data Infrastructure: Tracking Changes

It also allows for synchronization across systems:

A editor changes a record, which means the API needs updated, which means the website needs updated, and the search interfaces, and third-party systems...



```
content: "https://linked.data.gov/rdf/linked-act-api",
- class: [rdf:type],
+ label: "Subject Heading - Topical",
  id: "https://data.getty.edu/social/thesaurus/semantic-subject-topical",
  type: "Type",
},
},
content: "Art, American -- California -- 20th century",
content: "Krispy Krreme",
id: "https://data.getty.edu/research/collections/document/514879ac-162-552-917-8d727e",
- identifier_by: []
+ class: [rdf:type],
+ label: "generated slug",
  id: "https://data.getty.edu/social/thesaurus/temorary-value",
  type: "Type",
- {
+ label: "slug identifier",
  id: "https://data.getty.edu/social/thesaurus/Alum-identifier",
  type: "Type",
},
content: "1899-1900",
id: "https://data.getty.edu/research/collections/document/541867ac-162-552-917-8d727e",
type: "Identifier",
},
- part: []
+ {
  id: "https://data.getty.edu/research/collections/document/613322-602-542-4731",
  type: "linguistic-object"
},
},
type: "linguistic-object"
}
```



# Linked Data Infrastructure: ActivityStreams and Standards

The infrastructure uses the **W3C ActivityStream** standard and are implemented using the patterns from the **IIIF Change Discovery API**.

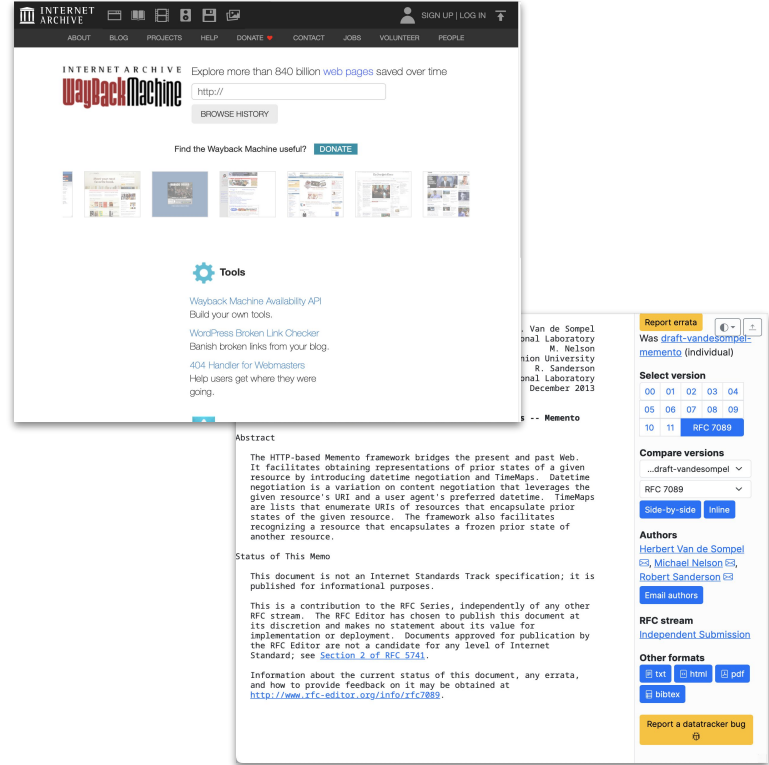
Using standards makes it easy to build integrations against changing data—both within our organization and for external aggregators.

The top screenshot shows the W3C Recommendation page for **Activity Streams 2.0**, dated 23 May 2017. It includes links for the latest published version, latest editor's draft, test suite, and implementation report. The bottom screenshot shows the **IIIF Change Discovery API 1.0** documentation page, featuring a navigation menu, a table of contents with sections like Introduction, Overview of IIIF Resource Discovery, and Status of this Document.

# Linked Data Infrastructure: ActivityStreams and Standards

For some kinds of data, it's also valuable to also be able to see **what has changed over time** for a given record.

To do so, our APIs also supports **Memento**, the standard underneath the Internet Archive.



## LLinked Data Infrastructure: ActivityStreams and Standards

This lets you automatically open older versions of the record—providing an **audit log** and the ability for scholars to understand how **knowledge changes over time**.

```
[
  - {
    from: "Sun, 24 Sep 2023 12:24:36 GMT",
    rel: "self",
    until: "Sun, 24 Sep 2023 13:38:37 GMT",
    uri: "http://localhost:5100/cidoc/-tm-/person/2"
  },
  - {
    rel: "original timegate",
    uri: "http://localhost:5100/cidoc/person/2"
  },
  - {
    datetime: "Sun, 24 Sep 2023 13:36:23 GMT",
    rel: "last memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/53ae2179-7327-460a-bd33-19b9618235d6"
  },
  - {
    datetime: "Sun, 24 Sep 2023 13:31:35 GMT",
    rel: "memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/aca563d9-e265-455c-9d17-a4385a6e13d6"
  },
  - {
    datetime: "Sun, 24 Sep 2023 13:10:54 GMT",
    rel: "memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/f7258356-fd31-4e23-9237-c7b1610bb562"
  },
  - {
    datetime: "Sun, 24 Sep 2023 13:10:23 GMT",
    rel: "memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/762dbcca-73e3-4a68-bb12-5bb0f574e282"
  },
  - {
    datetime: "Sun, 24 Sep 2023 12:59:42 GMT",
    rel: "memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/35ee9fb1-2ef5-459e-b89b-fa5fd4a742fe"
  },
  - {
    datetime: "Sun, 24 Sep 2023 12:59:12 GMT",
    rel: "memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/45f7d6a5-7a66-4e0b-a456-d5cc8e4dc44a"
  },
  - {
    datetime: "Sun, 24 Sep 2023 12:58:23 GMT",
    rel: "memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/3026f52b-ef14-402e-8913-f9b9442861ca"
  },
  - {
    datetime: "Sun, 24 Sep 2023 12:24:36 GMT",
    rel: "first memento",
    uri: "http://localhost:5100/cidoc/-VERSION-/f1623faa-ec39-491c-91a7-e5d5ed4d3bac"
  }
]
```

Cool Tech, Bro:

# Why Does this Matter?

## A Hard-won lesson:

No application that we've built **required** Linked Data.



## A Hard-won lesson:

No application that we've built **required** Linked Data.

Which, if you think about it, makes sense. Each application has a **specific, known context** with clear record boundaries.

## A Hard-won lesson:

Different users have different contexts and need different affordances.

A shared, graph-based data model **allows us to re-present the data** in a way that matches user's varying models of the world via multiple interfaces.

## **As Simple as possible:**

A shared data model also makes our developers more effective—eventually.

Building on top of web technology lets the engineering learning curve be gradual.

## As Simple as possible:

Standards are valuable for interoperability—but also because you don't have to **write all the documentation**.

Nobody wants to write it, but you can't work across institutions without it.

## **As Simple as possible:**

Minimize complexity in the data model.

Data is for computers—text is for humans. Resist the urge to show off.

You can always add complexity—you can **never take it away**.

## Disciplinary Misdeeds:

The hardest part of this will be **change management**.

Recontextualizing information across boundaries hides disciplinary labor—and digital innovations can conflict with pre-digital best practices.

## **Evangelize and collaborate.**

What makes cultural data interesting is *not* contained within any one institution.

It's shared across our entire, world-wide community. We should work together.

**Shared models and shared code make that easier.**

## What's Next: Provenance Index

Over the next several years, we'll be expanding our usage of this system:

This fall, we'll launch a new version of the Getty **Provenance Index**, adding in 22M records of transactions between art dealers.

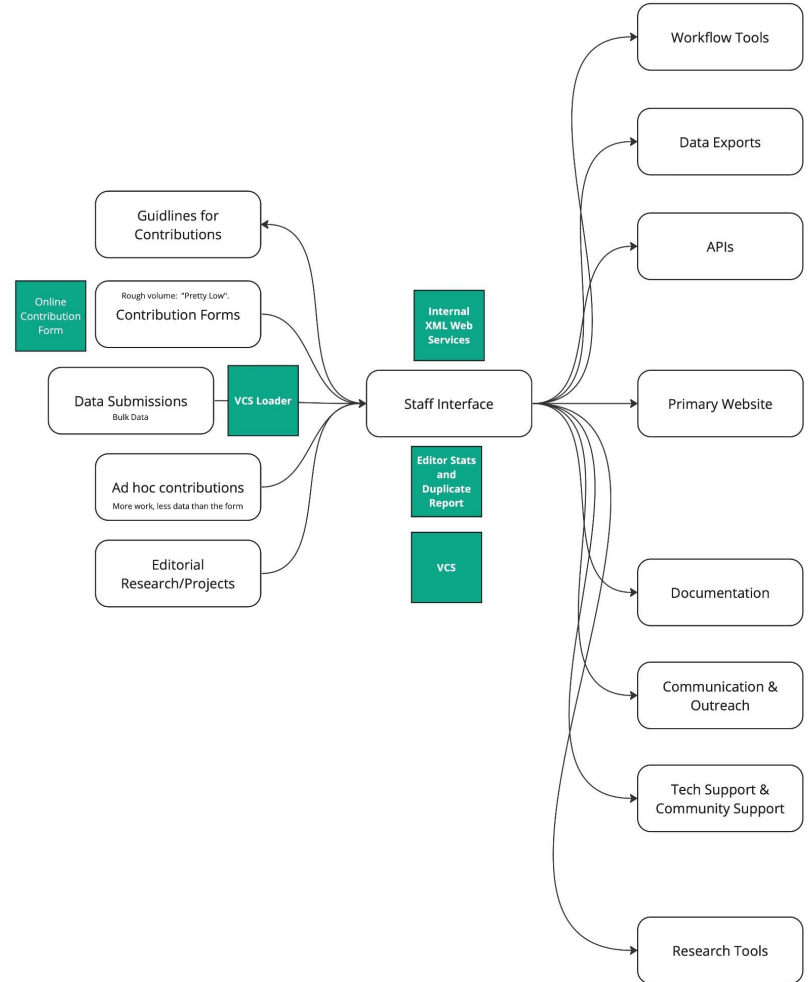
This **research-focused dataset** will allow new insights into collections around the world—and into the art market as a whole.



## What's Next: Getty Vocabularies

We're beginning to plan the next iteration of the **Getty Vocabularies** infrastructure:

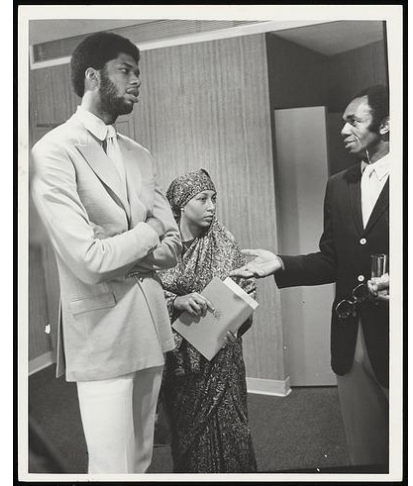
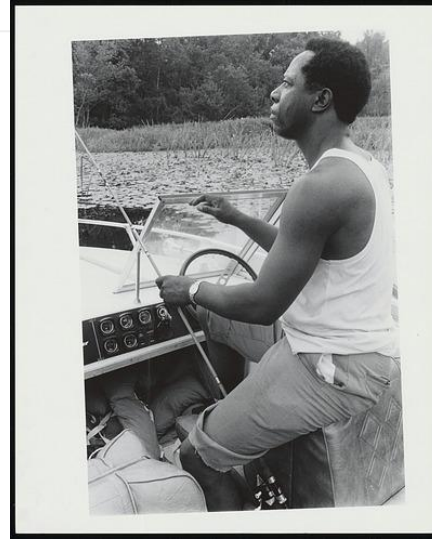
Working to understand how the multiple contexts of our **audiences** can be supported—and how new ways of working impact the platform.



## What's Next: Johnson Publishing Company Archives

And we're using the **platforms** and **standards** we've put in place to enable collaboration across the field:

Working with the Smithsonian to provide **joint access and discovery** for millions of images from the photo morgue of magazines such as Ebony and Jet.



## Why do we do Linked Data?

The value is not in the technologies or the ontologies we use.

## Why do we do Linked Data?

The value is in the **ecosystem**—information in varied context for different applications.

The value is in the **audience**—supporting user needs and conceptual models.

And it's in the **community**—allowing data and code to be used beyond the Getty.

## Why do we do Linked Data?

We do it for humans.

# Thank You.

Contact me at [dnewbury@getty.edu](mailto:dnewbury@getty.edu)

David Newbury

Assistant Director, Software and User Experience, Getty

CNI Spring Project Briefings, March 26, 2024